

553,617

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international(43) Date de la publication internationale  
4 novembre 2004 (04.11.2004)

PCT

(10) Numéro de publication internationale  
WO 2004/095783 A2(51) Classification internationale des brevets<sup>7</sup> : H04L 12/56(21) Numéro de la demande internationale :  
PCT/FR2004/000955

(22) Date de dépôt international : 16 avril 2004 (16.04.2004)

(25) Langue de dépôt : français

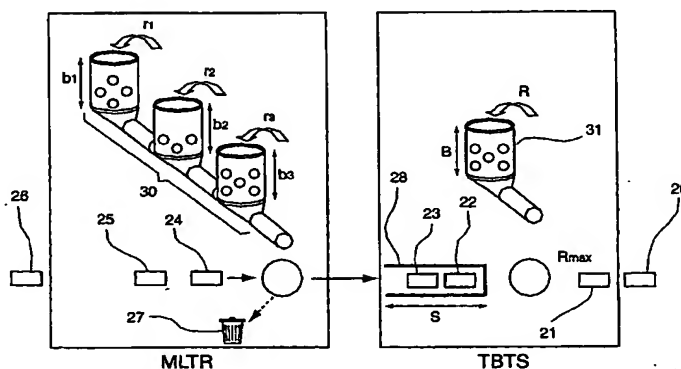
(26) Langue de publication : français

(30) Données relatives à la priorité :  
03/04903 18 avril 2003 (18.04.2003) FR(71) Déposant (pour tous les États désignés sauf US) :  
FRANCE TELECOM [FR/FR]; 6 Place d'Alleray,  
F-75015 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : BABON-  
NEAU, Gérard [FR/FR]; 8, Allée du marché, F-35235  
Thorigné fouillard (FR). LOUDHAIEF, Wissem [FR/FR];  
39Ter, Boulevard de la Liberté, F-35000 Rennes (FR).(74) Mandataire : GUENE, P.; Cabinet Vidon, 16B, rue de  
Jouanet-BP 90333, F-35703 Rennes Cedex 7 (FR).(81) États désignés (sauf indication contraire, pour tout titre de  
protection nationale disponible) : AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO,  
CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB,  
GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG,  
KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG,  
MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH,

[Suite sur la page suivante]

(54) Title: METHOD AND DEVICE FOR CONTROLLING DATA PACKET TRAFFIC AT THE INPUT OF A NETWORK, AND  
CORRESPONDING COMPUTER PROGRAM AND NETWORK EQUIPMENT(54) Titre : PROCEDE ET DISPOSITIF DE CONTROLE D'UN TRAFIC DE PAQUETS DE DONNEES EN ENTRÉE D'UN  
RESEAU, PROGRAMME D'ORDINATEUR ET EQUIPEMENT RESEAU CORRESPONDANTS

(57) Abstract: The invention relates to a method of controlling data packet traffic at the input of a network, whereby the traffic comprises N streams and/or substreams which are each associated with a priority level,  $N \geq 2$ , and each of the aforementioned packets is marked with the priority level associated with the stream or substream to which said packet belongs. The inventive method comprises a step employing a token bucket mechanism with N operating levels and N token buffers each containing a number of available tokens, the tokens of each of the N token buffers being used to process one of the N priority levels. Moreover, each of the packets is accepted or refused according to whether or not it is possible for same to be attributed tokens depending on the tokens available at least in the token buffer which is used to process the priority level of each packet. In one particular embodiment of the invention, the tokens from the N token buffers are shared between the N priority levels, and a packet with priority level i can be attributed tokens from a token buffer which is associated with priority level j, said level having less priority, when there are not sufficient tokens available in the i priority level token buffer.

(57) Abrégé : L'invention concerne un procede de controle d'un trafic de paquets de donnees en entree d'un reseau, le trafic comprenant N flux et/ou sous-flux associes chacun a un niveau de priorite,  $N \geq 2$ , chacun des paquets etant marque avec le niveau de priorite associe au flux ou sous-flux

[Suite sur la page suivante]

WO 2004/095783 A2



PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

- (84) États désignés (*sauf indication contraire, pour tout titre de protection régionale disponible*) : ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— sans rapport de recherche internationale, sera republiée dès réception de ce rapport

*En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.*

auquel appartient ledit paquet. Le procédé comprend une étape de mise en oeuvre d'un mécanisme de seuil à jetons à N niveaux de fonctionnement avec N mémoires-tampons de jetons contenant chacune un nombre de jetons disponibles, les jetons de chacune des N mémoires-tampons de jetons étant utilisés pour traiter l'un des N niveaux de priorité, chacun des paquets étant accepté ou refusé selon qu'il est possible ou non de lui attribuer des jetons en fonction des jetons disponibles au moins dans la mémoire-tampon de jetons utilisée pour traiter le niveau de priorité dudit paquet. Dans un mode de réalisation particulier, les jetons des N mémoires-tampons de jetons sont partagés entre les N niveaux de priorité, un paquet de niveau de priorité i pouvant se voir attribuer des jetons d'une mémoire-tampon de jetons associée à un niveau de priorité j moins prioritaire lorsque les jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité i ne sont pas suffisants.

**Procédé et dispositif de contrôle d'un trafic de paquets de données en entrée d'un réseau, programme d'ordinateur et équipement réseau correspondants.**

Le domaine de l'invention est celui des réseaux de communication, et notamment mais non exclusivement des réseaux de type IP ou équivalents.

5 Plus précisément, l'invention concerne un procédé de contrôle d'un trafic de paquets de données en entrée d'un réseau, dans le cas où le trafic comprend une pluralité de flux et/ou sous-flux associés chacun à un niveau de priorité, et où chacun des paquets est marqué avec le niveau de priorité associé au flux ou sous-flux auquel appartient ce paquet. En d'autres termes, l'invention concerne un mécanisme réseau permettant  
10 d'optimiser l'écoulement d'un trafic entrant sur un réseau.

L'invention a de nombreuses applications, telles que par exemple le contrôle de flux multimédia (par exemple des flux MPEG), seuls ou par agrégats, ou encore le contrôle d'un multiplex de flux de différentes natures (par exemple un flux vidéo/audio multiplexé avec un flux TCP, sur un accès ADSL).

15 On présente maintenant brièvement le contexte réseau dans lequel se situe la présente invention.

L'augmentation des performances des ordinateurs ainsi que les débits offerts par les nouvelles générations de réseaux ouvrent la voie à de nouveaux services basés sur les flux multimédias. De fait, le nombre d'informations audiovisuelles transmises sur les  
20 réseaux (par exemple de type IP (« Internet Protocol »)) croît sans cesse, et les algorithmes de compression (par exemple de type MPEG (« Moving Picture Experts Group »)) s'améliorent, offrant une qualité meilleure avec des débits plus faibles. Cependant, aujourd'hui le niveau de qualité n'est pas toujours acceptable. Si chaque maillon de la chaîne a les capacités intrinsèques pour fournir cette qualité, leur mise bout  
25 à bout et le partage des ressources réseau IP par de nombreux usagers aboutissent parfois à des résultats médiocres.

D'une manière générale, la transmission d'informations dans un réseau IP s'appuie sur la couche transport pour un contrôle de la qualité entre la source et les récepteurs. Cette couche, située entre le routage et les applications, est réalisée  
30 traditionnellement par le protocole TCP (« Transmission Control Protocol »). D'un point de vue application, TCP se charge de retransmettre les informations perdues ou mal

reçues, par un contrôle au niveau de la session. D'un point de vue réseau, certains paramètres du protocole permettent de déceler des possibles congestions, et d'adapter le débit de la source aux contraintes du réseau. L'objectif est alors de limiter le débit si le réseau ne peut pas tout écouler, et d'éviter d'envoyer des paquets qui seront perdus. De nombreuses études cherchent aujourd'hui à appliquer des mécanismes équivalents aux flux vidéo, avec la contrainte temps réel d'une adaptation dynamique des codeurs au débit disponible.

Mais en raison du temps important de réaction entre un ou plusieurs clients et la source vidéo, les protocoles temps réel et audiovisuels ne font actuellement que peu de traitements, et se limitent principalement au marquage du temps d'émission et à l'encapsulation des paquets de l'application pour leur routage dans la couche IP (par exemple RTP/UDP), à charge des applications de se débrouiller avec les données reçues.

L'évolution des réseaux offre la possibilité de gérer la « qualité de service » (QoS) dans les routeurs. Or, il est pertinent de remarquer que c'est le lieu où se produisent la plupart des pertes des réseaux IP, et des mécanismes sont mis en oeuvre à ce niveau pour traiter sélectivement les différents flux en vue d'assurer au mieux des objectifs de qualité. Les moyens déployés pour améliorer la qualité des flux transmis sont l'objet des travaux des groupes « IntServ » (pour « integrated services », c'est-à-dire « intégration de services ») et « DiffServ » (pour « differentiated services », c'est-à-dire « services différenciés ») à l'IETF (Internet Engineering Task Force) :

- « IntServ » définit des moyens pour réserver une ressource en débit garantie entre deux nœuds d'un réseau ;
- « DiffServ » définit des moyens pour contrôler dynamiquement le débit d'agrégats de flux en fonction de la charge du réseau.

En comparaison avec les solutions de bout en bout (analogues à TCP) les solutions localisées dans les routeurs ont plusieurs avantages :

- elles s'affranchissent de toute contrainte de session ;
- elles sont aussi adaptées au temps réel, car leur action dans les routeurs est immédiate sans nécessiter de retour d'informations en provenance des usagers ;

- elles sont donc naturellement adaptées à la diffusion sélective (« multicast »), car indépendantes du nombre d'utilisateurs alimentés par le flux et indépendantes de rapports en provenance d'un nombre variable d'utilisateurs.

Les traitements dans les routeurs s'appuient sur une distinction des paquets arrivant dans les routeurs supportant des mécanismes de « qualité de service » (QoS, pour « Quality of Service »).

Cette distinction existe naturellement dans les flux MPEG, car la compression des flux vidéo MPEG aboutit à une suite de données de natures différentes et non indépendantes. On distingue trois types d'images : I, P et B. Les images de type I (Intra), sont chacune codées sans faire référence à une autre image, la compression étant limitée à la réduction des redondances spatiales au sein de chaque image. Pour optimiser la quantité d'informations transmises, les codeurs MPEG exploitent le fait que deux images consécutives sont peu différentes en général. Une détection de mouvement ne considère que la partie qui vient de changer par rapport à l'image précédente pour obtenir une information de taille réduite codée dans une image de type P (Prédite). D'autres moyens permettent d'obtenir une information encore plus réduite en interpolant les mouvements entre deux images de type I ou P ; ces images sont alors de type B (Bidirectionnelle).

La taille des images P est beaucoup plus petite en général que celle des images I, et un codage avec peu d'images I permet d'obtenir, à débit équivalent, une qualité de décodage bien supérieure. Dans ces conditions, la perte d'une image n'est pas équivalente en fonction de la nature de l'information qu'elle contient. Cette structure de l'information doit conduire à considérer l'importance ou le poids de chaque information dans son traitement par le réseau.

Deux raisons justifient la conservation d'images I :

- la re-synchronisation périodique du flux en cas de pertes ;
- tout changement de scène car il n'est plus possible de s'appuyer sur les images précédentes pour un codage de mouvement.

Une autre manière de considérer ce poids de l'information est la découpe d'un flux MPEG4 en plusieurs niveaux hiérarchiques pour obtenir une qualité variable en fonction du contenu global reçu par l'utilisateur. Un niveau hiérarchique N doit s'appuyer

sur la présence des N-1 niveaux inférieurs pour apporter un complément de qualité. Un cas élémentaire est de considérer une vidéo constituée d'un flux de base (contenant des images I et P) et d'un flux de rehaussement (contenant des images P et B). Pour ce cas élémentaire, le flux de base dans sa globalité est considéré plus prioritaire que le flux de rehaussement.

La distinction de nature des images ou des flux dans le trafic MPEG est à exploiter dans les routeurs « DiffServ » pour traiter sélectivement les différentes informations d'un flux vidéo :

- soit par un marquage du champs TOS (« type of service », c'est-à-dire « type de service ») ou DSCP (« Differentiated Services Code Point », c'est-à-dire « valeur de marquage pour services différenciés ») des paquets par le serveur vidéo,
- soit par une classification effectuée par le routeur, ce qui aboutit également à un marquage des paquets IP.

Dans le présent document, on considère que le marquage des paquets est possible dans tous les cas. Les moyens de réaliser ce marquage sont considérés comme bien connus de l'homme du métier. Ils ne sont donc pas discutés en détail.

Quand une situation de congestion apparaît dans un routeur, les paquets reçus sont éliminés en fonction de la charge et de leur niveau de priorité.

Une caractéristique importante des données contenus dans ces paquets est leur variation importante en fonction du contenu de la scène. Or, les informations les plus critiques pour la restitution à l'utilisateur sont contenues dans les rafales les plus importantes, et le problème principal des réseaux IP de type « Best effort » (c'est-à-dire sans garantie) est leur difficulté à laisser passer les rafales en cas de congestion.

De ce fait, le simple marquage des flux élémentaires d'un flux MPEG n'est pas suffisant pour leur traitement optimal. En effet, les mécanismes « DiffServ » usuels sont prévus pour accepter des rafales que l'on trouve habituellement dans les réseaux IP, avec les applications qui sont majoritaires aujourd'hui : le transfert d'URL (« Uniform Resource Locator ») pour les applications WEB et les transferts de fichiers par FTP (« File Transport Protocol »). Toutes ces applications exploitent le protocole TCP dont

les mécanismes ont fait l'objet de nombreuses études afin d'obtenir une montée en charge progressive, et une adaptation du débit en fonction de la charge du réseau.

Or, les flux vidéo remettent en cause ce fonctionnement car leur comportement est qualifié d'agressif, dans la mesure où ils ignorent généralement l'état et la charge du réseau. De plus, la plupart des mécanismes introduits dans les réseaux IP ont tous pour objectif le lissage des flux pour favoriser un bon écoulement du trafic. Paradoxalement, les codeurs fournissent une qualité meilleure quand ils produisent des flux à débit variable, qui sont les plus maltraités dans les réseaux IP.

L'introduction de flux vidéo dans le réseau IP se heurte donc à trois contradictions :

- augmenter la qualité des codages, conduit pour un débit moyen défini, à produire des rafales de paquets quand la séquence codée l'exige ;
- les réseaux d'accès offrent une bande passante maximale limitée (y compris en ADSL (« Asymetric Digital Subscriber Line ») car les applications sont tentées d'exploiter un débit moyen proche du maximum pour procurer la meilleure qualité aux usagers ;
- les rafales sont les informations les plus importantes car elles correspondent à un changement de scène ou au moins à un changement important du contenu de l'image. Bien souvent, ces images sont du type I (Intra) dont la perte est très critique, car il devient impossible de reproduire les images suivantes, même si ces dernières sont correctement reçues.

On présente maintenant les techniques connues de conditionnement de trafic visant à réduire les congestions dans les réseaux.

On notera tout d'abord que les travaux sur l'application de la mise en forme pour les flux MPEG dans les réseaux IP en se basant sur UDP et RTP comme protocoles de transport sont très rares. La majorité des travaux concernent les réseaux ATM (« Asynchronous Transfer Mode ») et l'utilisation de TCP comme protocole de transport.

Les mécanismes de mise en forme et de conditionnement de trafic sont utilisés dans les réseaux IP à « qualité de service » (QoS). On peut notamment se reporter au document « Traffic Shaping for MPEG video transmission over next generation

internet » (lissage de trafic pour la transmission de vidéo MPEG sur l'internet de  
prochaine génération), M.F Alan, M. Atiquzzaman, M.A. Karim. Dans ce document, le  
lissage de trafic (TS) est utilisé pour assurer la conformité des flux MPEG au TSPEC  
(« Traffic Specifier » c'est-à-dire « spécification de trafic ») nécessaire à la réservation  
de ressources dans les réseaux « IntServ ». Pour les réseaux « DiffServ », le traitement  
se fait par agrégat de flux et donc sans trop se soucier des applications. Les algorithmes  
de lissage de trafic (TS) sont par contre largement utilisés au niveau du codage dans le  
but de contrôler le débit des codeurs. Ceci reste insuffisant pour maîtriser les flux au  
niveau réseau.

L'utilisation de la mise en forme, par lissage de trafic (TS, pour « Traffic  
Shaping ») ou contrôle des règles de trafic (ou TP, pour « Traffic Policing »), pour  
réduire la congestion dans le réseau peut améliorer d'une manière significative le niveau  
de « qualité de service » (QoS) que le réseau est capable de délivrer aux applications.

Le lissage (TS) lisse les rafales en « bufferisant » (c'est-à-dire en mettant en  
mémoire-tampon) les paquets concernés par l'excès de rafales dans les équipements de  
frontière du réseau. Il peut réduire la congestion à des niveaux acceptables, surtout que  
les algorithmes d'ordonnancement (« scheduler ») tels que CBQ (« Class Based  
Queuing » c'est-à-dire « gestion de files d'attente basées par classes de services ») ou  
bien PQ (« Priority Queuing » c'est-à-dire « gestion de files d'attentes par priorités ») ne  
sont pas capables de le faire. Utilisés seuls, ces mécanismes propagent les rafales dans le  
réseau.

Comme le lissage (TS), le contrôle des règles de trafic (TP) limite le débit du  
trafic au débit configuré. Mais au lieu de « bufferiser » les paquets comme le lissage  
(TS), les paquets non-conformes sont soit rejetés, soit remarqués pour diminuer leurs  
priorités. Le contrôle des règles de trafic (TP) ne lisse pas par conséquent le trafic mais  
il n'introduit pas de délai de « bufferisation » non plus.

Dans la majorité des architectures qui supportent la « qualité de service » (QoS),  
un contrat de service existe entre le fournisseur du service réseau (NSP, pour « Network  
Service Provider ») et le fournisseur de l'application (ASP, pour « Application Service  
Provider »). Dans les réseaux ATM, ce contrat est appelé « Contrat de trafic ». Dans les  
réseaux « DiffServ », les aspects de contractualisation sont traités dans le SLA



(« Service Level Agreement ») et plus précisément dans le SLS (« Service Level Specification »).

On connaît également le document suivant : RFC2475, « An Architecture for Differentiated Services » (une architecture pour des services différenciés), Décembre 1998. Il précise que les sources de trafic peuvent assurer les tâches de classification et de conditionnement de trafic. En effet, le trafic peut être marqué avant de quitter le domaine source. Ce marquage initial est appelé « Initial Marking » ou « Pre-marking ». L'avantage du marquage initial est que les préférences et les besoins des applications peuvent être mieux pris en compte pour décider quels paquets doivent recevoir un meilleur traitement dans le réseau.

La prévention contre la surcharge d'une classe de service donnée n'est pas une tâche facile dans les réseaux « DiffServ ». En plus, en cas de surcharge dans une classe de service, il faut noter que tous les flux de cette classe de service souffrent d'une dégradation de la « qualité de service » (QoS).

En plus, plusieurs mécanismes utilisés dans la mise en œuvre de « DiffServ » fonctionnent d'une façon moins efficace en présence des rafales. Le mécanisme RED (« Random Early Drop »), par exemple, est plus efficace lorsqu'il est appliqué à un trafic lissé, sinon ce sont les flux lissés qui sont pénalisés alors que les flux en rafales ne subissent pas une amélioration significative.

Les flux MPEG sont caractérisés par le fait qu'ils présentent des rafales (nature « bursty ») et par leur sensibilité aux pertes de paquets. Ces pertes causent une dégradation de la qualité subjective, mais il est très difficile de prévoir le niveau de cette dégradation suite à des pertes. Cette dégradation est fortement liée à la nature des informations transportées par les paquets perdus. Les mécanismes de correction d'erreurs sont utilisés lors du décodage pour palier aux pertes.

Gérer les flux MPEG dans le réseau, ou même dans les routeurs d'extrémité (ER, pour « Edge Router »), est une tâche compliquée. Le fournisseur du service réseau (NSP) n'est pas obligé de traiter différemment les flux MPEG. En plus, l'agrégat de trafic résultant de plusieurs flux audio-visuels est généralement difficile à décrire :

- le processus d'arrivée de paquets est auto-similaire ;
- grande variation des données transportées par les paquets ;

- la dynamique des protocoles.

Une solution consiste à marquer les paquets et à leur attribuer le niveau de « qualité de service » (QoS) approprié avant qu'ils quittent le domaine du fournisseur de service sur Internet (ISP, pour « Internet Service Provider »). La passerelle d'accès au média (MAG, pour « Media Access Gateway ») est par exemple responsable de cette tâche. Cette passerelle MAG gère le trafic selon le SLS spécifié. Cette approche facilite la négociation de SLA/SLS pour les services en mode continu (« Streaming ») et impose au client un profil de trafic bien particulier.

Parmi les techniques actuelles, la plus répandue pour contrôler le trafic est le mécanisme WRED (« Weighted Random Early Drop ») qui consiste en une perte de paquets aléatoire et différente en fonction du marquage des paquets. Ce mécanisme se base sur un taux de remplissage moyen de la file d'attente d'émission sur un lien d'un réseau. Mais cette technique introduit un caractère aléatoire des rejets de paquets, et le taux de remplissage de la file d'attente n'est pas optimisée. Selon les tailles des rafales et leur fréquence, ces rejets peuvent se produire pour un remplissage faible ou un remplissage très élevé de la file d'attente. Ce qui conduit d'une part à une sous utilisation de la file d'attente, et d'autre part à la réservation de taille mémoire conséquente pour la réalisation de cette file. Ce problème existe pour tout type d'application, et il est encore plus réel pour les flux audiovisuels en raison de leurs rafales importantes.

Pour détailler ce point, il est fondamental de remarquer tout d'abord que les rafales des flux vidéo sont imprévisibles, autant en taille qu'en durée, tout en garantissant un débit moyen sur une période de l'ordre d'une seconde. Cette fenêtre de calcul du débit moyen est beaucoup trop grande pour obtenir des tailles raisonnables des files d'attente associées. De fait, les routeurs réagissent en calculant les moyennes de remplissage sur une dizaine de paquets reçus, alors que certaines rafales peuvent largement dépasser les 20 paquets.

Ainsi, une succession de rafales peut aboutir à des rejets par saturation de la capacité de la file d'attente, inhibant le fonctionnement normal du mécanisme. A l'opposé, quand un trafic faible survient après une suite de rafales, la valeur moyenne

peut rester temporairement anormalement élevée et des paquets sont rejetés alors que la file est pratiquement vide.

Le mécanisme WRED n'est donc pas adapté au contrôle de flux caractérisés par une valeur moyenne sur une durée fixe.

5 L'invention a notamment pour objectif de pallier ces inconvénients de l'état de la technique et offrir une solution optimale en cas de congestion du réseau.

10 Plus précisément, l'un des objectifs de la présente invention est de fournir un procédé et un dispositif de contrôle de trafic permettant de contrôler des rafales et de lisser le trafic sur un ensemble de flux et/ou sous-flux associés à des niveaux de priorités.

15 En d'autres termes, un objectif de la présente invention est de fournir un procédé et un dispositif de contrôle de trafic permettant de protéger les informations importantes des rafales, afin d'apporter une solution à l'antinomie entre l'optimisation d'un flux (par exemple un flux vidéo) contenant des rafales et le lissage des rafales pour un transport de qualité dans le réseau.

20 L'invention a également pour objectif de fournir de tels procédé et dispositif qui soient simples à mettre en œuvre et peu coûteux.

Un autre objectif de l'invention est de fournir de tels procédé et dispositif permettant de proposer efficacement des contrats de trafic (SLA/SLS) entre opérateurs  
25 réseau et fournisseurs de services.

Ces différents objectifs, ainsi que d'autres qui apparaîtront par la suite, sont atteints selon l'invention à l'aide d'un procédé de contrôle d'un trafic de paquets de données en entrée d'un réseau, le trafic comprenant N flux et/ou sous-flux associés chacun à un niveau de priorité,  $N \geq 2$ , chacun des paquets étant marqué avec le niveau  
30 de priorité associé au flux ou sous-flux auquel appartient ledit paquet, ledit procédé comprenant une étape de mise en œuvre d'un mécanisme de seau à jetons à N niveaux de fonctionnement avec N mémoires-tampons de jetons contenant chacune un nombre de jetons disponibles, les jetons de chacune des N mémoires-tampons de jetons étant utilisés pour traiter l'un des N niveaux de priorité, chacun des paquets étant accepté ou refusé selon qu'il est possible ou non de lui attribuer des jetons en fonction

des jetons disponibles au moins dans la mémoire-tampon de jetons utilisée pour traiter le niveau de priorité dudit paquet.

Le principe général de l'invention consiste donc à utiliser un seau à jetons multi-niveau (MLTB, pour « Multi Layer Token Bucket ») pour rejeter les paquets en dehors d'un profil requis et caractérisé par les N niveaux de fonctionnement du seau à jetons multi-niveau. Chaque paquet subit un traitement selon un marquage correspondant à son niveau de priorité. Les paquets acceptés sont placés dans une file d'attente.

Le seau à jetons multi-niveau permet de traiter sélectivement et conjointement plusieurs niveaux de priorités de flux. Il est bien adapté à la caractérisation du trafic entre la taille des rafales en entrée et l'écoulement du trafic en sortie. On sait qu'en cas de congestion du réseau, il est illusoire de chercher de la bande passante disponible. Le réglage des paramètres du seau à jetons multi-niveau assure une relation entre les niveaux de priorité pour équilibrer les contraintes de fonctionnement entre les débits par niveaux de priorité et les rafales acceptables par le seau à jeton selon les contraintes des applications transportées. La caractérisation des paramètres des N jeux de paramètres du seau à jetons multi-niveau permet de nombreuses solutions et peut s'adapter à peu près à tous les cas de fonctionnement :

- un cas extrême est le comportement avec N jeux de paramètres indépendants agissant comme des seaux à jetons indépendants ;
- un autre cas extrême est la possibilité pour le niveau le plus prioritaire de prendre tous les jetons et d'aboutir à un rejet des paquets de tous les autres niveaux ;
- entre ces deux extrêmes, toutes les configurations sont possibles, autorisant plus ou moins de rafales ou plus ou moins de débit réservé par niveau.

L'invention permet donc de traiter des rafales sur les niveaux les plus prioritaires, du fait qu'il existe pour chaque niveau de priorité une réserve disponible pour prévenir toute arrivée soudaine d'un ensemble de données qui ne doivent pas être rejetées.

On rappelle qu'un seau à jeton habituel ne possède qu'un seul niveau de fonctionnement (il dispose d'un unique jeu de paramètres) et traite donc tous les paquets

indistinctement. En cas de congestion du réseau, les paquets sont donc rejetés indépendamment de leur niveau de priorité.

On notera que la présente invention est totalement compatible avec les flux IP « unicast » (à destinataire unique) et « multicast » (diffusés sélectivement).

5 On notera également que la présente invention permet de transmettre dans une même classe de service plusieurs groupes de flux avec des priorités différentes. En particulier, l'invention permet de fournir un traitement adapté à un ou un groupe de flux vidéo (IPB ou hiérarchique) conformément à un profil de trafic contractualisé (SLS) par des valeurs caractéristiques de type seau à jeton. En effet, les paramètres facilement  
10 mesurables et adaptables d'un seau à jetons multi-niveau (MLTB) sont un moyen efficace de proposer des contrats de trafic (SLA/SLS) entre opérateurs réseau et fournisseurs de services. La présence d'informations prioritaires conduit à la spécification de ce seau. Les multiples déclinaisons de ce seau sont un moyen d'offrir des classes de services adaptées aux exigences des clients. Quelles que soient les  
15 applications, le profil de trafic fait intervenir les principaux éléments de caractérisation d'un flux dans un réseau : le débit et le délai. L'invention est donc un moyen de définir un contrat avec un compromis négocié entre le débit, la taille des rafales, et le délai de transmission.

Dans une première application avantageuse de l'invention, le trafic comprend N  
20 sous-flux correspondant chacun à un des N niveaux hiérarchiques d'un flux hiérarchique ou d'un agrégat de flux hiérarchiques.

Il s'agit par exemple d'un flux hiérarchique audio/vidéo comprenant les sous-flux suivants : un sous-flux audio, un sous-flux vidéo de base et un sous-flux vidéo de réhaussement.

25 Dans une deuxième application avantageuse de l'invention, le trafic comprend N sous-flux correspondant chacun à un des N types d'images d'un flux multimédia ou d'un agrégat de flux multimédia.

Il s'agit par exemple d'un flux vidéo MPEG comprenant trois sous-flux correspondant aux trois types d'images I, P et B.

30 Dans une troisième application avantageuse de l'invention, le trafic comprend N flux correspondant chacun à un des flux d'un multiplex d'au moins deux flux.

Il s'agit par exemple d'un flux vidéo/audio multiplexé avec un flux TCP, sur un accès ADSL.

Avantageusement, le trafic comprend N flux et/ou sous-flux appartenant à une même classe de service.

5 Ainsi, l'invention, grâce au seau à jetons multi-niveau, permet de transmettre dans une même classe de service plusieurs flux et/ou sous-flux avec des priorités différentes. On peut distinguer plusieurs classes de service, telles que par exemple la classe « streaming » (contenant des flux audio et vidéo, la classe « TCP prioritaire », la classe « Best Effort de réseau IP », ...) et utiliser un seau à jeton multi-niveau pour  
10 chaque classe. Chaque classe de service peut être définie par marquage des paquets, par l'adresse IP source ou destination des paquets, par le protocole utilisé pour les paquets, etc.

Préférentiellement, les paquets refusés sont jetés.

Les paquets refusés sont ceux qui ne sont pas conformes au profil de trafic défini  
15 par les paramètres des N niveaux de fonctionnement du seau à jetons multi-niveau.

Une autre option, moins performante mais qui rentre néanmoins dans le cadre de la présente invention, est de transmettre les paquets refusés après les avoir à nouveau marqués avec un niveau de priorité plus faible. Cette option présente toutefois l'inconvénient d'augmenter la probabilité de rejet de paquets dans les nœuds  
20 congestionnés du réseau.

Avantageusement, le réseau est de type IP ou équivalent.

Selon une caractéristique avantageuse, chacun des N niveaux de fonctionnement du mécanisme de seau à jetons est géré par un régulateur  $b_i(r_i, bm_i)$ ,  $i \in \{1 \text{ à } N\}$ , avec :

- $r_i$  le débit nominal du régulateur ;
- 25 -  $bm_i$  la taille maximale de la mémoire-tampon de jetons du régulateur ;
- $b_i(t)$  la valeur instantanée du remplissage de la mémoire-tampon de jetons du régulateur.

De façon préférentielle, les jetons des N mémoires-tampons de jetons sont partagés entre les N niveaux de priorité, un paquet de niveau de priorité i pouvant se voir  
30 attribuer des jetons d'une mémoire-tampon de jetons associée à un niveau de priorité j

moins prioritaire lorsque les jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité *i* ne sont pas suffisants.

De cette façon, on favorise les rafales des niveaux les plus prioritaires.

Avantageusement, pour chaque niveau de priorité hormis le niveau de priorité le plus prioritaire, on garantit une quantité de jetons réservés exclusivement aux paquets possédant ledit niveau de priorité.

Ainsi, on garantit une ressource minimum pour chaque niveau de priorité, sans exclure un usage partiellement partagé des ressources (c'est-à-dire des jetons des différents niveaux de fonctionnement du seau à jetons multi-niveau).

Dans un premier mode de réalisation particulier de l'invention, l'attribution de jetons à un paquet de niveau de priorité *i* est effectuée selon un mode discontinu par paquet et consiste à attribuer :

- soit des jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité *i* ;
- soit des jetons disponibles dans une mémoire-tampon de jetons de niveau de priorité *j* moins prioritaire, lorsque les jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité *i* ne sont pas suffisants.

Ainsi, dans le mode discontinu, un paquet ne peut se servir des ressources (jetons) que d'un seul niveau de fonctionnement du seau à jetons multi-niveau.

Dans un second mode de réalisation particulier de l'invention, l'attribution de jetons à un paquet de niveau de priorité *i* est effectuée selon un mode continu par bit et consiste à attribuer :

- des jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité *i* ;
- et en complément des jetons disponibles dans au moins une mémoire-tampon de jetons de niveau de priorité *j* moins prioritaire, lorsque les jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité *i* ne sont pas suffisants.

En d'autres termes, dans le mode continu, un paquet peut se servir des ressources (jetons) de plusieurs niveaux de fonctionnement du seau à jetons multi-niveau à la fois.

De façon avantageuse, les paquets acceptés par le mécanisme de seau à jetons à *N* niveaux de fonctionnement sont placés dans une file d'attente. Ledit procédé comprend en outre une étape de mise en œuvre d'un mécanisme de seau à jetons à un

seul niveau de fonctionnement avec une unique mémoire-tampon de jetons, de façon à prendre les paquets contenus dans la file d'attente et les envoyer sur le réseau en exécutant un lissage du trafic par limitation du débit instantané à une valeur acceptable par le réseau.

5           Le seau à jetons à un seul niveau de fonctionnement (TBTS, pour « Token Bucket Traffic Shaper ») permet donc de limiter les débits crêtes émis par l'équipement réseau supportant l'invention. Il réalise une temporisation des rafales quand elles dépassent le débit toléré dans le réseau.

10           On utilise par exemple une zone mémoire, dont l'entrée est gérée par le seau multi-niveau (contrôle des rafales) et dont la sortie est gérée par le seau à un seul niveau (lissage de trafic).

L'invention concerne également un programme d'ordinateur comprenant des instructions de code de programme pour l'exécution des étapes du procédé tel que précité, lorsque ledit programme est exécuté sur un ordinateur.

15           L'invention concerne aussi un dispositif de contrôle d'un trafic de paquets de données en entrée d'un réseau, le trafic comprenant N flux et/ou sous-flux associés chacun à un niveau de priorité,  $N \geq 2$ , chacun des paquets étant marqué avec le niveau de priorité associé au flux ou sous-flux auquel appartient ledit paquet,

20           ledit dispositif comprenant des moyens de mise en œuvre d'un mécanisme de seau à jetons à N niveaux de fonctionnement avec N mémoires-tampons de jetons contenant chacune un nombre de jetons disponibles, les jetons de chacune des N mémoires-tampons de jetons étant utilisés pour traiter l'un des N niveaux de priorité, chacun des paquets étant accepté ou refusé selon qu'il est possible ou non de lui attribuer des jetons en fonction des jetons disponibles au moins dans la mémoire-tampon  
25           de jetons utilisée pour traiter le niveau de priorité dudit paquet.

De façon préférentielle, ledit procédé comprend des moyens de partage des jetons des N mémoires-tampons de jetons entre les N niveaux de priorité, un paquet de niveau de priorité i pouvant se voir attribuer des jetons d'une mémoire-tampon de jetons associée à un niveau de priorité j moins prioritaire lorsque les jetons disponibles dans la  
30           mémoire-tampon de jetons de niveau de priorité i ne sont pas suffisants.



Avantageusement, pour chaque niveau de priorité hormis le niveau de priorité le plus prioritaire, lesdits moyens de partage comprennent des moyens de garantie d'une quantité de jetons réservés exclusivement aux paquets possédant ledit niveau de priorité.

L'invention concerne encore un équipement réseau comprenant un dispositif de contrôle tel que précité, ledit équipement réseau appartenant au groupe comprenant :

- des équipements réseau situés entre un réseau d'un fournisseur d'application ou de service et un réseau d'un fournisseur d'un service réseau, constituant ledit réseau en entrée duquel est effectué le contrôle d'un trafic de paquets de données ;
- des routeurs compris dans des nœuds d'un réseau d'un fournisseur d'un service réseau, constituant ledit réseau en entrée duquel est effectué le contrôle d'un trafic de paquets de données

D'autres caractéristiques et avantages de l'invention apparaîtront à la lecture de la description suivante d'un mode de réalisation préférentiel de l'invention, donné à titre d'exemple indicatif et non limitatif, et des dessins annexés, dans lesquels :

- la figure 1 présente un exemple d'architecture de réseaux dans laquelle peut être mis en œuvre le procédé de contrôle de trafic selon l'invention ;
- la figure 2 illustre un mode de réalisation particulier du procédé selon l'invention, mettant en œuvre deux fonctions, basées sur l'utilisation respectivement d'un seau à jetons multi-niveau et un seau à jetons à niveau unique ;
- la figure 3 illustre un régulateur gérant un niveau de fonctionnement du seau à jetons multi-niveau illustré sur la figure 2, ainsi que l'utilisation dans ce régulateur d'un paramètre  $K_j$  garantissant un minimum de ressource pour le niveau de priorité associé à ce régulateur ;
- la figure 4 illustre un exemple d'attribution de jetons dans le cas où le fonctionnement du seau à jetons multi-niveau est décrit avec des mémoires-tampons (buffers) indépendantes ;
- la figure 5 illustre un exemple d'attribution de jetons dans le cas où le fonctionnement du seau à jetons multi-niveau est décrit avec des mémoires-tampons (buffers) corrélées.

L'invention concerne donc un procédé de contrôle d'un trafic de paquets de données en entrée d'un réseau. Le trafic est du type comprenant N flux et/ou sous-flux associés chacun à un niveau de priorité, avec  $N \geq 2$ . Chacun des paquets est marqué avec le niveau de priorité associé au flux ou sous-flux auquel il appartient.

5 A titre d'exemple, l'invention permet de transmettre en priorité les informations essentielles d'un flux vidéo, ou de plusieurs flux vidéo regroupés dans un agrégat. Selon la nature du flux, cette distinction est par exemple possible soit par les images de type IBP (voir définition ci-dessus), soit par les n couches d'un flux hiérarchique. Si dans le premier cas le débit moyen des images I reste faible devant le débit global, dans le  
10 second cas l'information la plus importante et à protéger au maximum est définie par la fraction du débit global occupé par la couche de base, et pouvant représenter jusqu'à 50% du flux. En général, la couche de base est la seule à contenir les informations de références contenues dans les images I.

Dans un agrégat, toutes les données de même niveau de priorité subissent le  
15 même traitement. Par exemple, tous les flux de base ou toutes les images I sont traités comme un seul flux de débit plus important qu'une vidéo seule.

On présente maintenant, en relation avec la figure 1, un exemple d'architecture de réseaux dans laquelle peut être mis en œuvre le procédé de contrôle de trafic selon l'invention.

20 Dans cet exemple, on considère le cas de fournisseurs de service (ISP, pour « Internet Service provider ») offrant un service de transmission de vidéo en mode continu (« streaming »).

L'exemple d'architecture (simplifiée) illustrée sur la figure 1 comprend :

- 25 – un réseau IP de type « DiffServ » 1, géré par un opérateur réseau (aussi appelé fournisseur du service réseau, ou NSP (pour « Network service Provider »)) et comprenant une pluralité de routeurs  $2_1$  à  $2_5$  compris dans des nœuds du réseau IP ;
- deux réseaux de fournisseurs de service « streaming » 3 et 4, l'un comprenant deux serveurs 5 et 6, et l'autre un serveur 7. Chaque réseau de fournisseurs de  
30 service 3, 4 comprend en outre un équipement d'extrémité 8, 9 connecté à l'un des routeurs du réseau IP, appelé routeur d'extrémité (« edge routeur »)

2<sub>1</sub>, qui est chargé d'envoyer les données d'un serveur vers une artère principale d'un cœur de réseau IP ;

- deux terminaux clients 10 et 11, chacun connecté au réseau IP 1 par exemple via un lien ADSL 12, 13.

5 On suppose qu'il existe deux contrats de trafic (symbolisés par les flèches référencées SLA1 et SLA2 respectivement sur la figure 1) : l'un entre l'opérateur du réseau IP 1 et le fournisseur de service dont le réseau est référencé 3, et l'autre entre l'opérateur du réseau IP 1 et le fournisseur de service dont le réseau est référencé 4.

10 Le procédé selon l'invention est mis en œuvre dans un équipement réseau formant un conditionneur de trafic pour l'entrée dans le réseau IP 1. Cet équipement réseau peut être situé entre le réseau du fournisseur du service 3, 4 et le réseau IP 1 :

- soit dans l'un des équipements d'extrémité 8, 9 des réseaux de fournisseurs de service 3 et 4 ;
- soit dans le routeur d'extrémité 2<sub>1</sub> du réseau IP 1.

15 L'équipement réseau implémentant le procédé selon l'invention peut également être tout routeur placé au niveau d'un point de congestion dans le réseau (en particulier pour tout accès à un lien ADSL).

20 On présente maintenant, en relation avec la figure 2, un mode de réalisation particulier du procédé selon l'invention. Il permet de protéger les informations importantes des rafales et apporter des solutions à l'antinomie entre l'optimisation d'un flux vidéo contenant des rafales et le lissage des rafales pour un transport de qualité dans le réseau.

25 Dans ce mode de réalisation particulier du procédé selon l'invention, le mécanisme d'ensemble forme un « conditionneur de trafic décentralisé multi-couche » appelé «MLDTC» (pour « Multi-Layers Decentralized Traffic Conditioner »). Il est constitué de deux fonctions exécutées séquentiellement :

- la première est appelée « MLTR » (pour « Multi-Layers Traffic Regulator ») du fait qu'elle forme un « régulateur de trafic multi-couche ». Elle assure le contrôle de rafales par niveau hiérarchique. Elle est basée sur la mise en œuvre d'un seau à jetons à N niveaux de fonctionnement (référéncé 30), avec N buffers (mémoires-tampons) de jetons virtuels. Chaque buffer de jetons

correspond à un niveau hiérarchique. Chaque niveau de fonctionnement est défini par un jeu de paramètres (voir discussion détaillée ci-après). La fonction MLTR assure l'attribution de jetons en fonction de la priorité du paquet entrant et de la ressource disponible dans le buffer de jetons correspondant. Dans l'exemple de la figure 2, on a  $N = 3$  et la représentation de trois seaux indique uniquement la distinction des trois jeux de paramètres (il n'y a en réalité qu'un seul seau à trois niveaux de fonctionnement) ;

- la seconde est appelée « TBTS » (pour « Token Bucket Traffic Shaper ») du fait qu'elle est basée sur un algorithme de type seau à jetons à niveau unique (« Token Bucket ») (référéncé 31) et exécute un lissage du trafic pour un écoulement régulier sur le réseau en limitant le débit instantané à une valeur acceptable par le réseau. La mémoire tampon de la fonction TBTS est de capacité raisonnable pour garantir un coût moindre des systèmes et un temps de transfert limité dans la fourniture de l'information à l'utilisateur. La contrepartie de la taille finie est la limitation des rafales acceptées en entrée du réseau.

La fonction MLTR est adaptée à l'acceptation différenciée des rafales en fonction du niveau de priorité. Elle s'applique aux paquets IP (référéncés 20 à 26 sur la figure 2) dont le champ DSCP (« Differentiated Services Code point ») est marqué soit par la source, soit par un système de classification. Chaque niveau de priorité  $i$  dispose de ses propres paramètres, pour un tampon unique, commun aux 3 niveaux, mais avec des niveaux d'acceptation des paquets de données variables. Ce mécanisme répond aux attentes d'un signal vidéo, car il ne traite pas uniformément toutes les informations. En effet, on rappelle qu'un codage MPEG aboutit à plusieurs niveaux d'informations qu'il convient de traiter différemment selon leur importance. L'objectif est de favoriser les rafales d'un niveau  $i$  par rapport à un niveau  $j$ , dont les informations sont moins importantes ( $j > i$ ). Les paquets relatifs au niveau  $i$  ont alors la priorité pour exploiter les ressources disponibles.

Grâce à la fonction TBTS, le même débit moyen est obtenu pour des tailles de rafales différentes, et tenant compte de l'intervalle les séparant. Un tel mécanisme est

bien adapté aux rafales des flux vidéo, qui sont variables, en raison de la grande diversité des contenus transportés (sports, actualités, paysages, ...).

On présente maintenant de façon détaillée un exemple de réalisation de la fonction MLTR, qui assure la régulation et le contrôle de rafales. Dans la suite de la description, on étudie à titre d'exemple le cas où cette fonction MLTR est basée sur un seau à jetons à trois niveaux de fonctionnement ( $N = 3$ ).

Comme illustré sur la figure 3, chaque niveau de fonctionnement  $i$  du seau à jetons à trois niveaux de fonctionnement 30 est géré par un régulateur  $b_i(r_i, bm_i)$ ,  $i \in \{1, 2, 3\}$ , avec :

- $r_i$  le débit nominal de ce régulateur ;
- $bm_i$  la taille maximale du buffer de jetons de ce régulateur ;
- $b_i(t)$  la valeur instantanée du remplissage du buffer de jetons de ce régulateur.

La taille du buffer de jetons ( $b_i$ ) impose une taille maximale pour les rafales pour chaque niveau. La tolérance d'un régulateur de niveau  $i$  vis-à-vis des rafales de grande taille dépend de la taille du buffer de jetons  $b_i$  et du débit  $r_i$ .

Les paquets conformes (c'est-à-dire ceux à qui il a pu être attribué des jetons par le seau multi-niveau 30) sont placés dans un buffer de paquets à émettre 28, qui forme un moyen de gestion d'une file d'attente. Si un nombre insuffisant de jetons est disponible dans le seau multi-niveau, les paquets sont alors considérés comme non-conformes et sont éliminés. Ils sont jetés dans la poubelle référencée 27. Une autre option est de transmettre ces paquets avec une priorité plus faible. Cependant, le nouveau marquage des paquets pour leur attribuer une priorité plus faible augmente la probabilité de rejet de paquets dans les nœuds congestionnés.

Pour favoriser les rafales des niveaux les plus prioritaires, les trois buffers de jetons,  $b_1$ ,  $b_2$  et  $b_3$  sont partagés entre les différents niveaux. En d'autres termes, un paquet de niveau de priorité  $i$  peut se voir attribuer des jetons d'un buffer de jetons associé à un niveau de priorité  $j$  moins prioritaire ( $j > i$ ) lorsque les jetons disponibles dans le buffer de jetons de niveau de priorité  $i$  ne sont pas suffisants.

Cependant, à cause de ce principe d'emprunt, les paquets de niveau  $i$  risquent d'utiliser toutes les ressources en jetons des niveaux moins prioritaires  $j$ . Ceci peut empêcher de garantir le débit nominal au niveau  $j$  au profit du niveau  $i$ . Comme illustré

sur la figure 3, la solution est de limiter cet emprunt par un paramètre  $K_j$ , propre au niveau  $j$ , et dont l'objectif est de protéger ce dernier des niveaux plus prioritaires en garantissant une quantité de ressources (jetons) réservées exclusivement aux paquets de niveau  $j$ . Par conséquent, le paramètre  $K_j$  garantit un seuil de remplissage des buffers de jetons au-dessous duquel les paquets d'un niveau plus prioritaire ne peuvent plus se servir. Le paramètre  $K_j$  du régulateur de niveau  $j$  est illustré sur la figure 3.

La valeur de  $K_j$  doit être choisie telle que  $bm_j \geq K_j \geq MTU$ .

Ceci revient à garantir pour les paquets de niveau  $j$  le débit nominal et des rafales limitées par  $K_j$ . Le cas particulier où  $K_j$  est adapté à la taille utile d'un paquet sur le réseau (MTU (« Maximum Transfer Unit », c'est-à-dire « taille maximale d'une unité élémentaire de transfert ») pour un réseau IP) correspond à une garantie du débit nominal  $r_j$  pour le niveau  $j$  mais sans garantie de rafales. Les ressources en jetons pourraient être utilisées par les niveaux plus prioritaires.

Les trois jeux de paramètres définissant les trois niveaux de fonctionnement du seau à jetons multi-niveau peuvent être décrits :

- soit par une représentation de trois buffers indépendants, avec un usage partagé des ressources entre les trois niveaux ;
- soit par une représentation cumulée des ressources faisant mieux ressortir les priorités entre les niveaux et les interdépendances de fonctionnement provenant du partage des ressources.

Toutefois, ces représentations sont identiques dans leur fonctionnement.

On présente maintenant, en relation avec la figure 4, une description du fonctionnement du seau à jetons multi-niveau avec des buffers (mémoires-tampons) indépendants.

Le rechargement des buffers de jetons est calculé pour chacun des trois niveaux au moment de l'arrivée d'un nouveau paquet, en fonction du temps le séparant du paquet précédent.

$$\begin{cases} b_1(t+T) = \min[(b_1(t) + r_1.T), bm_1] \\ b_2(t+T) = \min[(b_2(t) + r_2.T), bm_2] \\ b_3(t+T) = \min[(b_3(t) + r_3.T), bm_3] \end{cases} \quad (\text{Équation 1})$$

La prise en compte du paquet arrivant de niveau  $i$  et de taille  $P_i$  est obtenue par l'algorithme suivant : (mode paquets)

$$\left\{ \begin{array}{l} si(P_1) \left\{ \begin{array}{l} si(P_1 \leq b_1) \Rightarrow b_1 = b_1 - P_1; envoi; \\ autre : si(P_1 \leq (b_2 - K_2)) \Rightarrow b_2 = b_2 - P_1; envoi; \\ autre : si(P_1 \leq (b_3 - K_3)) \Rightarrow b_3 = b_3 - P_1; envoi; \\ autre : perte; \end{array} \right. \\ \\ si(P_2) \left\{ \begin{array}{l} si(P_2 \leq b_2) \Rightarrow b_2 = b_2 - P_2; envoi; \\ autre : si(P_2 \leq (b_3 - K_3)) \Rightarrow b_3 = b_3 - P_2; envoi; \\ autre : perte; \end{array} \right. \\ \\ si(P_3) \left\{ \begin{array}{l} si(P_3 \leq b_3) \Rightarrow b_3 = b_3 - P_3; envoi; \\ autre : perte; \end{array} \right. \end{array} \right. \quad (\text{Équation 2})$$

5 La figure 4 est une représentation graphique de cet algorithme d'attribution de jetons dans le cas où le fonctionnement du seau à jetons multi-niveau est décrit avec des buffers indépendants. Elle montre l'usage possible des ressources  $b_1$ ,  $b_2$  et  $b_3$  pour chacun des niveaux. Les traits épais (référéncés 41, 42 et 43) montrent le remplissage des buffers de jetons (c'est-à-dire le nombre de jetons disponibles) pour chacun des niveaux, et les flèches les ressources (c'est-à-dire les jetons) prises par les paquets entrants ( $P_1$ ,  $P_2$  ou  $P_3$ ) selon leur priorité.

On présente maintenant, en relation avec la figure 5, une description du fonctionnement du seau à jetons multi-niveau avec des buffers corrélés.

Le fonctionnement avec trois régulateurs indépendants  $b_i(r_i, bm_i)$ ,  $i \in \{1, 2, 3\}$ , est équivalent au fonctionnement avec trois régulateurs dépendants  $B_i(R_i, BM_i)$  avec :

- 15
- $B_i(t)$  la valeur instantanée du remplissage du buffer de jetons virtuel relatif aux paquets de niveau  $i$  ;
  - $BM_i$  la taille maximale du buffer de jetons virtuel relatif aux paquets de niveau  $i$ .

Les tailles de buffers de jetons par niveau sont obtenus par :

$$\begin{cases} B_3 = b_3 \\ B_2 = b_2 + b_3 \\ B_1 = b_1 + b_2 + b_3 \end{cases} \quad (\text{Équation 3})$$

Ce mode implique que :  $B_1 \geq B_2 \geq B_3$

Le rechargement des buffers de jetons peut alors s'exprimer sous une forme équivalente à Equation 1 par :

$$\begin{cases} B_1(t+T) = \min[(b_1(t) + r_1.T), bm_1] + \min[(b_2(t) + r_2.T), bm_2] + \min[(b_3(t) + r_3.T), bm_3] \\ B_2(t+T) = \min[(b_2(t) + r_2.T), bm_2] + \min[(b_3(t) + r_3.T), bm_3] \\ B_3(t+T) = \min[(b_3(t) + r_3.T), bm_3] \end{cases} \quad (\text{Équation 4})$$

Si on suppose que  $\min[(b_i(t) + r_i.T), bm_i] = b_i(t) + r_i.T$ , (Equation 4) devient :

$$\begin{cases} B_1(t+T) = b_1(t) + b_2(t) + b_3(t) + (r_1 + r_2 + r_3)T \\ B_2(t+T) = b_2(t) + b_3(t) + (r_2 + r_3)T \\ B_3(t+T) = b_3(t) + r_3.T \end{cases}$$

La figure 5 est une représentation graphique de l'algorithme d'attribution de jetons dans le cas où le fonctionnement du seau à jetons multi-niveau est décrit avec des buffers corrélés. Elle montre l'usage possible des ressources  $B_1$ ,  $B_2$  et  $B_3$  pour chacun des niveaux. De même que sur la figure 4, les traits épais (référéncés 51, 52 et 53) montrent le remplissage des buffers de jetons (c'est-à-dire le nombre de jetons disponibles) pour chacun des niveaux, et les flèches les ressources (c'est-à-dire les jetons) prises par les paquets entrants (P1, P2 ou P3) selon leur priorité.

Si les équations du rechargement traduisent exactement le même comportement pour les deux approches, le choix de l'attribution des jetons aux paquets à l'arrivée peut se réaliser par deux alternatives :

- en mode discontinu (paquets): un paquet ne peut se servir que d'un seul régulateur ;
- en mode continu (bits): un paquet peut se servir de plusieurs régulateurs à la fois.



Selon le fonctionnement en mode discontinu (paquets), la prise en compte de l'arrivée du paquet est obtenue par l'algorithme suivant :

$$\left. \begin{array}{l}
 si(P_1) \left\{ \begin{array}{l}
 si(P_1 \leq B_1 - B_2) \Rightarrow B_1 = B_1 - P_1; envoi; \\
 autre : si(P_1 \leq ((B_2 - B_3) - K_2)) \Rightarrow \begin{cases} B_2 = B_2 - P_1; envoi; \\ B_1 = B_1 - P_1 \end{cases} \\
 autre : si(P_1 \leq (B_3 - K_3)) \Rightarrow \begin{cases} B_3 = B_3 - P_1 \\ B_2 = B_2 - P_1; envoi; \\ B_1 = B_1 - P_1 \end{cases} \\
 autre : perte;
 \end{array} \right. \\
 si(P_2) \left\{ \begin{array}{l}
 si(P_2 \leq B_2 - B_3) \Rightarrow \begin{cases} B_2 = B_2 - P_2; envoi; \\ B_1 = B_1 - P_2 \end{cases} \\
 autre : si(P_2 \leq (B_3 - K_3)) \Rightarrow \begin{cases} B_3 = B_3 - P_2 \\ B_2 = B_2 - P_2; envoi; \\ B_1 = B_1 - P_2 \end{cases} \\
 autre : perte;
 \end{array} \right. \\
 si(P_3) \left\{ \begin{array}{l}
 si(P_3 \leq B_3) \Rightarrow \begin{cases} B_3 = B_3 - P_3 \\ B_2 = B_2 - P_3; envoi; \\ B_1 = B_1 - P_3 \end{cases} \\
 autre : perte;
 \end{array} \right.
 \end{array} \right. \quad (Équation 5)$$

5 Il est à noter que (Equation 2) et (Equation 5) sont équivalentes.

Dans ce mode de fonctionnement discontinu (paquets), un paquet est servi par les ressources du niveau correspondant au paquet (niveau i) ou d'un niveau moins prioritaire (niveau j). Il est à noter que ce fonctionnement n'est pas optimal, car un paquet entrant est rejeté même si la somme des ressources disponibles réparties sur les 3

10 niveaux est suffisante.

Selon le fonctionnement en mode continu (bits), la prise en compte de l'arrivée du paquet est obtenue par l'algorithme suivant :

$$\left\{ \begin{array}{l}
 \text{si}(P_1) \left\{ \begin{array}{l}
 \text{si}(P_1 \leq b_1) \Rightarrow b_1 = b_1 - P; \text{envoi}; \\
 \text{autre: si}((P_1 - b_1) \leq (b_2 - K_2)) \Rightarrow \begin{cases} b_2 = b_2 - (P_1 - b_1) \\ b_1 = 0 \end{cases}; \text{envoi}; \\
 \text{autre: si}((P_1 - (b_1 + (b_2 - K_2))) \leq (b_3 - K_3)) \Rightarrow \begin{cases} b_3 = b_3 - (P_1 - (b_1 + (b_2 - K_2))) \\ b_2 = K_2 \\ b_1 = 0 \end{cases}; \text{envoi}; \\
 \text{autre: perte};
 \end{array} \right. \\
 \text{si}(P_2) \left\{ \begin{array}{l}
 \text{si}(P_2 \leq b_2) \Rightarrow b_2 = b_2 - P_2; \text{envoi}; \\
 \text{autre: si}(P_2 \leq (b_3 - K_3)) \Rightarrow \begin{cases} b_3 = b_3 - P_2 \\ b_2 = 0 \end{cases}; \text{envoi}; \\
 \text{autre: perte};
 \end{array} \right. \\
 \text{si}(P_3) \left\{ \begin{array}{l}
 \text{si}(P_3 \leq b_3) \Rightarrow b_3 = b_3 - P_3; \text{envoi}; \\
 \text{autre: perte};
 \end{array} \right.
 \end{array} \right.$$

A priori ce mode de fonctionnement continu (bits) est plus optimal que le mode discontinu (paquets). En effet, il permet une utilisation maximale de ressources propres du niveau avant de demander des ressources à un niveau moins prioritaire.

5 On présente maintenant de façon détaillée un exemple de réalisation de la fonction TBTS qui assure la mise en forme du trafic.

Le seau à jetons à niveau unique 31, sur lequel est basé cette fonction TBTS, est par exemple défini par les paramètres suivants :

- 10 – R, le débit moyen ;
- $R_{\max}$ , le débit maximal à la sortie ;
- B(t), la valeur instantanée du remplissage du buffer de jetons ( $B_{\max}$  est la taille maximale du buffer de jetons) ;
- S, la valeur instantanée du remplissage du buffer de paquets 28 ( $S_{\max}$  est la taille maximale du buffer de paquets 28) ;

- $T$ , le temps séparant l'arrivée de deux paquets consécutifs ;
- $\tau$ , la durée d'une rafale de paquets.

Les paquets conformes (c'est-à-dire ceux à qui il a pu être attribué des jetons par le seau à niveau unique 31) sont transmis alors que les paquets non-conformes sont retardés dans le buffer de paquets 28 pour les contenir dans l'enveloppe de trafic ciblé. La taille du buffer de paquets 28 est à manipuler avec précaution car elle peut induire un délai supplémentaire, qui doit rester raisonnable.

Le fonctionnement du seau à niveau unique 31 (et donc de la fonction TBTS) peut être décrit par les équations suivantes :

$$B(t+T) = \min[(B(t) + R.T), B_{\max}]$$

$$R_{\max}.\tau \leq B_{\max} + R.\tau \Rightarrow \tau \leq \frac{B_{\max}}{R_{\max} - R}$$

La prise en compte du paquet  $P_i$  est obtenue par l'algorithme suivant :

$$si(P_i) \begin{cases} si(P_i \leq B) \Rightarrow B = B - P_i; envoi; \\ autre : bufferisation; \end{cases}$$

Si on considère  $r_{\max}$ , le débit maximal à la sortie de la fonction MLTR, l'évolution du remplissage du buffer de paquets 28 entre l'arrivée de deux paquets consécutifs peut être décrit par l'équation suivante :

$$S(t+T) \leq S(t) - B_{\max} + (r_{\max} - R)T$$

$$et \ S(t) - R.T - B_{\max} \leq S(t+T) \leq S(t) + (r_{\max} - R)T$$

## REVENDICATIONS

1. Procédé de contrôle d'un trafic de paquets de données en entrée d'un réseau, le trafic comprenant N flux et/ou sous-flux associés chacun à un niveau de priorité,  $N \geq 2$ , chacun des paquets étant marqué avec le niveau de priorité associé au flux ou sous-flux auquel appartient ledit paquet,

caractérisé en ce qu'il comprend une étape de mise en œuvre d'un mécanisme de seau à jetons à N niveaux de fonctionnement avec N mémoires-tampons de jetons contenant chacune un nombre de jetons disponibles, les jetons de chacune des N mémoires-tampons de jetons étant utilisés pour traiter l'un des N niveaux de priorité, chacun des paquets étant accepté ou refusé selon qu'il est possible ou non de lui attribuer des jetons en fonction des jetons disponibles au moins dans la mémoire-tampon de jetons utilisée pour traiter le niveau de priorité dudit paquet.

2. Procédé selon la revendication 1, caractérisé en ce que le trafic comprend N sous-flux correspondant chacun à un des N niveaux hiérarchiques d'un flux hiérarchique ou d'un agrégat de flux hiérarchiques.

3. Procédé selon la revendication 1, caractérisé en ce que le trafic comprend N sous-flux correspondant chacun à un des N types d'images d'un flux multimédia ou d'un agrégat de flux multimédia.

4. Procédé selon la revendication 1, caractérisé en ce que le trafic comprend N flux correspondant chacun à un des flux d'un multiplex d'au moins deux flux.

5. Procédé selon l'une quelconque des revendications 1 à 4, caractérisé en ce que le trafic comprend N flux et/ou sous-flux appartenant à une même classe de service.

6. Procédé selon l'une quelconque des revendications 1 à 5, caractérisé en ce que les paquets refusés sont jetés.

7. Procédé selon l'une quelconque des revendications 1 à 6, caractérisé en ce que le réseau est de type IP ou équivalent.

8. Procédé selon l'une quelconque des revendications 1 à 7, caractérisé en ce que chacun des N niveaux de fonctionnement du mécanisme de seau à jetons est géré par un régulateur  $b_i(r_i, bm_i)$ ,  $i \in \{1 \text{ à } N\}$ , avec :

- $r_i$  le débit nominal du régulateur ;
- $bm_i$  la taille maximale de la mémoire-tampon de jetons du régulateur ;

- $b_i(t)$  la valeur instantanée du remplissage de la mémoire-tampon de jetons du régulateur.

9. Procédé selon l'une quelconque des revendications 1 à 8, caractérisé en ce que les jetons des N mémoires-tampons de jetons sont partagés entre les N niveaux de priorité, un paquet de niveau de priorité  $i$  pouvant se voir attribuer des jetons d'une mémoire-tampon de jetons associée à un niveau de priorité  $j$  moins prioritaire lorsque les jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité  $i$  ne sont pas suffisants.

10. Procédé selon la revendication 9, caractérisé en ce que, pour chaque niveau de priorité hormis le niveau de priorité le plus prioritaire, on garantit une quantité de jetons ( $K_j$ ) réservés exclusivement aux paquets possédant ledit niveau de priorité.

11. Procédé selon l'une quelconque des revendications 9 et 10, caractérisé en ce que l'attribution de jetons à un paquet de niveau de priorité  $i$  est effectuée selon un mode discontinu par paquet et consiste à attribuer :

- soit des jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité  $i$  ;
- soit des jetons disponibles dans une mémoire-tampon de jetons de niveau de priorité  $j$  moins prioritaire, lorsque les jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité  $i$  ne sont pas suffisants.

12. Procédé selon l'une quelconque des revendications 9 et 10, caractérisé en ce que l'attribution de jetons à un paquet de niveau de priorité  $i$  est effectuée selon un mode continu par bit et consiste à attribuer :

- des jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité  $i$  ;
- et en complément des jetons disponibles dans au moins une mémoire-tampon de jetons de niveau de priorité  $j$  moins prioritaire, lorsque les jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité  $i$  ne sont pas suffisants.

13. Procédé selon l'une quelconque des revendications 1 à 12, caractérisé en ce que les paquets acceptés par le mécanisme de seau à jetons à N niveaux de fonctionnement sont placés dans une file d'attente,

et en ce que ledit procédé comprend en outre une étape de mise en œuvre d'un mécanisme de seau à jetons à un seul niveau de fonctionnement avec une unique

mémoire-tampon de jetons, de façon à prendre les paquets contenus dans la file d'attente et les envoyer sur le réseau en exécutant un lissage du trafic par limitation du débit instantané à une valeur acceptable par le réseau.

14. Programme d'ordinateur, caractérisé en ce qu'il comprend des instructions de code de programme pour l'exécution des étapes du procédé selon l'une quelconque des revendications 1 à 3, lorsque ledit programme est exécuté sur un ordinateur.

15. Dispositif de contrôle d'un trafic de paquets de données en entrée d'un réseau, le trafic comprenant N flux et/ou sous-flux associés chacun à un niveau de priorité,  $N \geq 2$ , chacun des paquets étant marqué avec le niveau de priorité associé au flux ou sous-flux auquel appartient ledit paquet,

caractérisé en ce qu'il comprend des moyens de mise en œuvre d'un mécanisme de seau à jetons à N niveaux de fonctionnement avec N mémoires-tampons de jetons contenant chacune un nombre de jetons disponibles, les jetons de chacune des N mémoires-tampons de jetons étant utilisés pour traiter l'un des N niveaux de priorité, chacun des paquets étant accepté ou refusé selon qu'il est possible ou non de lui attribuer des jetons en fonction des jetons disponibles au moins dans la mémoire-tampon de jetons utilisée pour traiter le niveau de priorité dudit paquet.

16. Dispositif selon la revendication 15, caractérisé en ce qu'il comprend des moyens de partage des jetons des N mémoires-tampons de jetons entre les N niveaux de priorité, un paquet de niveau de priorité i pouvant se voir attribuer des jetons d'une mémoire-tampon de jetons associée à un niveau de priorité j moins prioritaire lorsque les jetons disponibles dans la mémoire-tampon de jetons de niveau de priorité i ne sont pas suffisants.

17. Dispositif selon la revendication 16, caractérisé en ce que, pour chaque niveau de priorité hormis le niveau de priorité le plus prioritaire, lesdits moyens de partage comprennent des moyens de garantie d'une quantité de jetons ( $K_j$ ) réservés exclusivement aux paquets possédant ledit niveau de priorité.

18. Equipement réseau comprenant un dispositif de contrôle selon l'une quelconque des revendications 15 à 17, caractérisé en ce que ledit équipement réseau appartient au groupe comprenant :

- des équipements réseau situés entre un réseau d'un fournisseur d'application ou de service et un réseau d'un fournisseur d'un service réseau, constituant ledit réseau en entrée duquel est effectué le contrôle d'un trafic de paquets de données ;
- 5 - des routeurs compris dans des nœuds d'un réseau d'un fournisseur d'un service réseau, constituant ledit réseau en entrée duquel est effectué le contrôle d'un trafic de paquets de données.

1/3

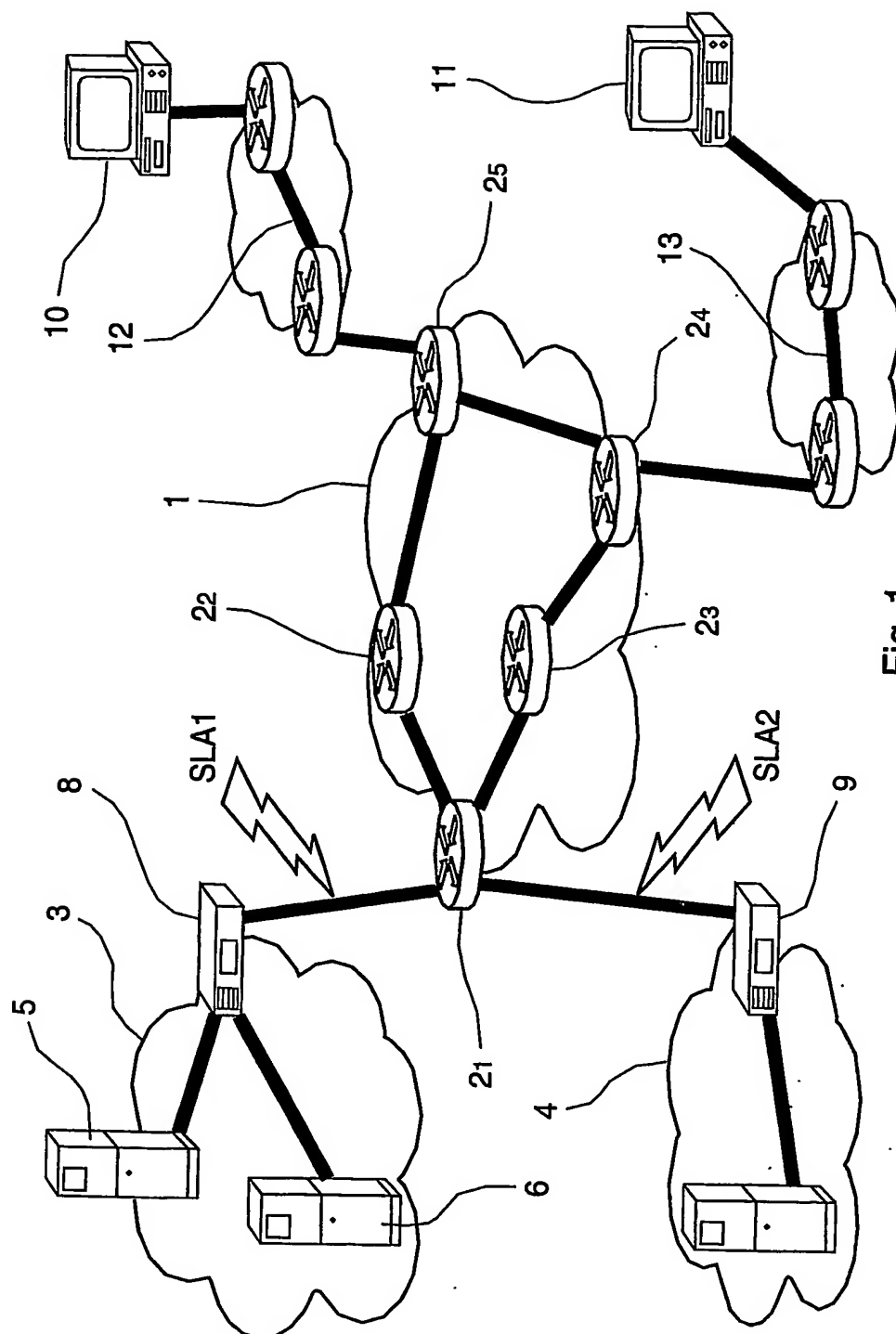
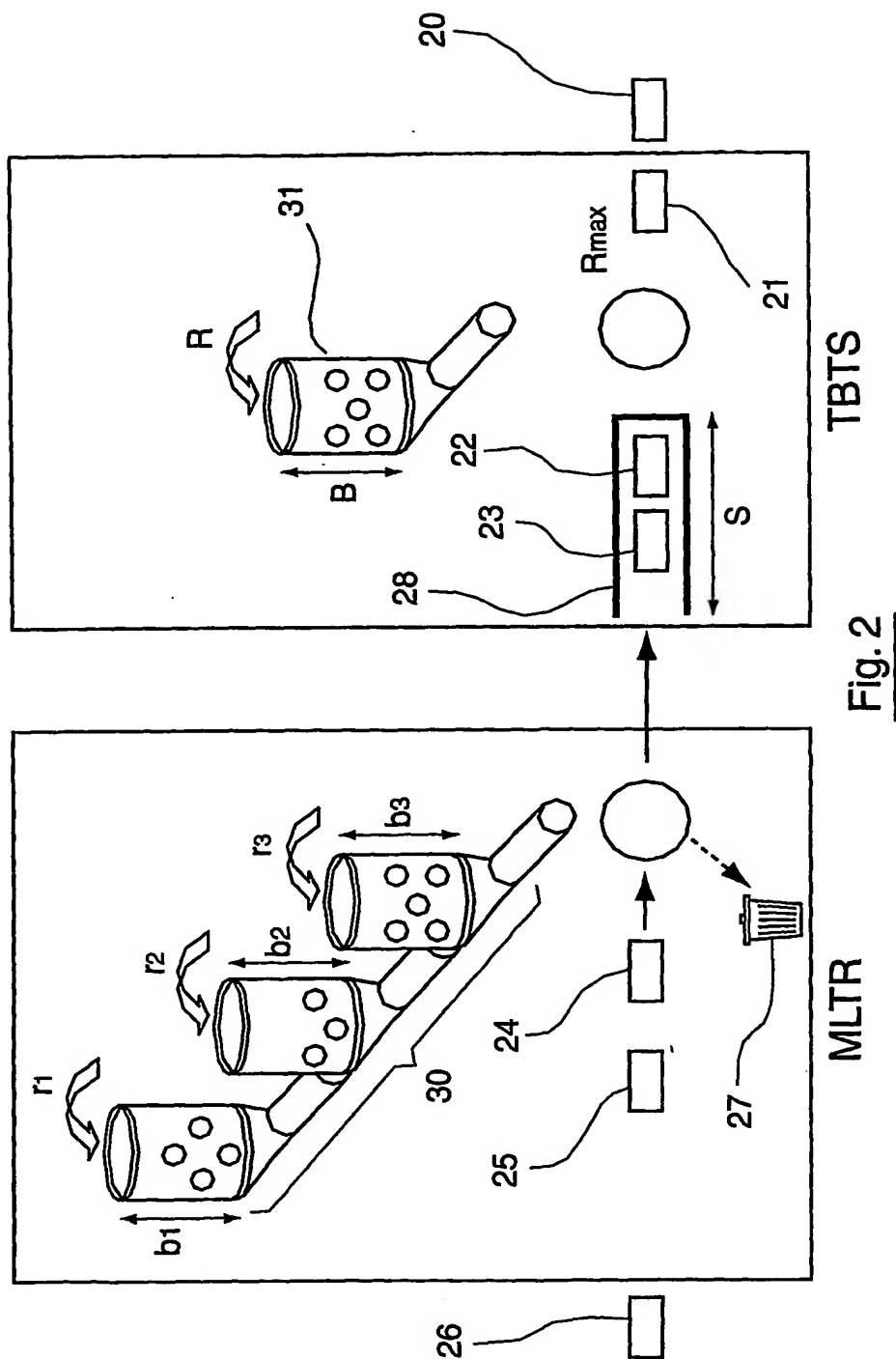


Fig. 1



2/3



3/3

Fig. 3

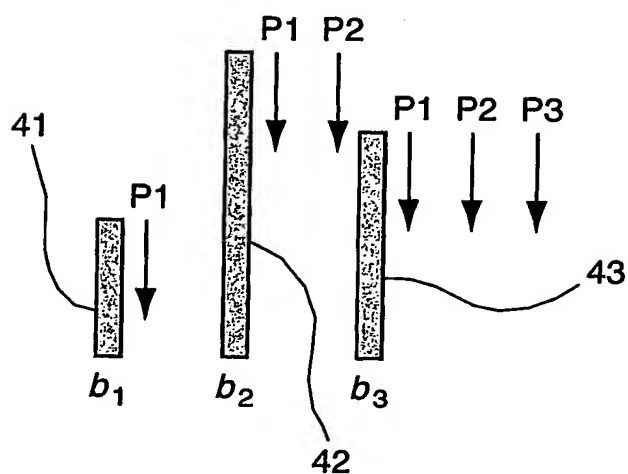
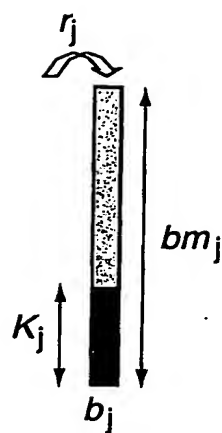


Fig. 4

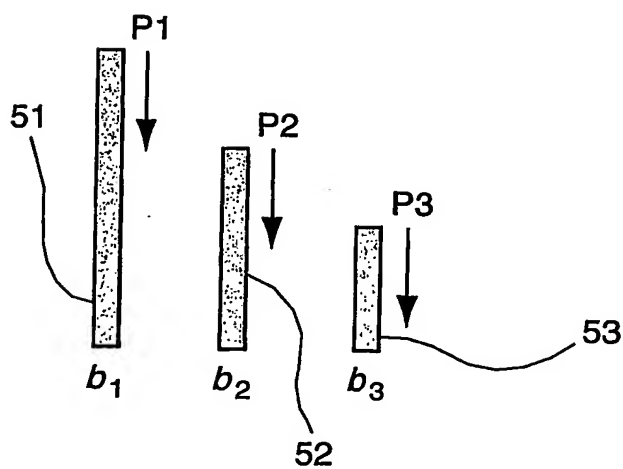


Fig. 5